日本国特許庁 JAPAN PATENT OFFICE

25.07.03

REC'D 12 SEP 2003

WIRE ROT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application:

2003年 5月12日

出 願 番 号 Application Number:

特願2003-133171

[ST. 10/C]:

[JP2003-133171]

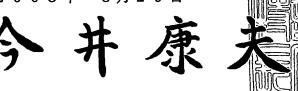
出 願 人
Applicant(s):

松下電器産業株式会社



SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

特許庁長官 Commissioner, Japan Patent Office 2003年 8月29日



Best Available Copy

【書類名】

特許願

【整理番号】

2903250015

【提出日】

平成15年 5月12日

【あて先】

特許庁長官殿

【国際特許分類】

G06K 9/46

【発明者】

【住所又は居所】

神奈川県横浜市港北区網島東四丁目3番1号 パナソニ

ックモバイルコミュニケーションズ株式会社内

【氏名】

藤田 幸宏

【発明者】

【住所又は居所】

神奈川県横浜市港北区網島東四丁目3番1号 パナソニ

ックモバイルコミュニケーションズ株式会社内

【氏名】

沢村 恒治

【発明者】

【住所又は居所】

神奈川県横浜市港北区綱島東四丁目3番1号 パナソニ

ックモバイルコミュニケーションズ株式会社内

【氏名】

伊藤 功将

【特許出願人】

【識別番号】

000005821

【氏名又は名称】

松下電器產業株式会社

【代理人】

【識別番号】

100105647

【弁理士】

【氏名又は名称】 小栗 昌平

【電話番号】

03-5561-3990

【選任した代理人】

【識別番号】

100105474

【弁理士】

【氏名又は名称】 本多 弘徳

【電話番号】

03-5561-3990

【選任した代理人】

【識別番号】

100108589

【弁理士】

【氏名又は名称】 市川 利光

【電話番号】

03-5561-3990

【選任した代理人】

【識別番号】 100115107

【弁理士】

【氏名又は名称】 高松 猛

【電話番号】 03-5561-3990

【選任した代理人】

【識別番号】

100090343

【弁理士】

【氏名又は名称】 栗宇 百合子

【電話番号】

03-5561-3990

【先の出願に基づく優先権主張】

【出願番号】

特願2002-230312

【出願日】

平成14年 8月 7日

【手数料の表示】

【予納台帳番号】 092740

【納付金額】

21,000円

【提出物件の目録】

【物件名】

明細書 1

【物件名】

図面 1

【物件名】

要約書 1

【包括委任状番号】 0002926

【プルーフの要否】 要



【書類名】 明細書

【発明の名称】 文字認識処理装置及び文字認識処理方法、並びに携帯端末装置 【特許請求の範囲】

【請求項1】 文字認識用に撮影された認識対象となる文字画像の画像データを取り込む画像取り込み手段と、

前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、

前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、

前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手 段と、

前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手 段と、

を備えた文字認識処理装置。

【請求項2】 請求項1記載の文字認識処理装置であって、

前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた 複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の 画像データを分割して取り込み、

前記レイアウト解析手段は、前記取り込んだ複数の文字画像の画像データごと に、前記カーソル位置情報を照合して文字の配置を解析するものである文字認識 処理装置。

【請求項3】 請求項1または2に記載の文字認識処理装置を備えた携帯端末装置。

【請求項4】 文字認識用の認識対象となる文字画像を撮影する撮影手段と

前記撮影された文字画像の画像データを取り込む画像取り込み手段と、

前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、

前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合し



て文字の配置を解析するレイアウト解析手段と、

前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手 段と、

前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手 段と、

を備えた携帯端末装置。

【請求項5】 請求項4記載の携帯端末装置であって、

前記撮影手段による撮影時に、前記文字画像と重ねて文字認識用文字枠を表示する文字枠表示手段を備える携帯端末装置。

【請求項6】 請求項4記載の携帯端末装置であって、

前記文字認識手段による認識結果の文字情報を表示する認識文字表示手段を備える携帯端末装置。

【請求項7】 請求項6記載の携帯端末装置であって、

前記認識文字表示手段は、前記認識結果の文字情報を所定の文字単位ごとに個別に選択可能に表示するものである携帯端末装置。

【請求項8】 請求項4記載の携帯端末装置であって、

前記文字認識手段による認識結果の文字情報を保存する認識文字保存手段を備 える携帯端末装置。

【請求項9】 請求項8記載の携帯端末装置であって、

前記認識文字保存手段は、前記文字情報を認識文字格納領域に保存するものである携帯端末装置。

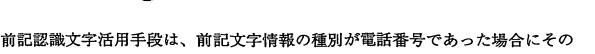
【請求項10】 請求項8記載の携帯端末装置であって、

前記認識文字保存手段は、前記文字情報の種別が電話番号、メールアドレス、 URL (Uniform Resource Locator) のいずれかであった場合に、その文字情報 の種別に応じたデータベースに登録するものである携帯端末装置。

【請求項11】 請求項4記載の携帯端末装置であって、

前記文字認識手段による認識結果の文字情報を当該文字情報の種別に応じて活用する認識文字活用手段を備える携帯端末装置。

【請求項12】 請求項11記載の携帯端末装置であって、



【請求項13】 請求項11記載の携帯端末装置であって、

電話番号への発信画面を表示するものである携帯端末装置。

前記認識文字活用手段は、前記文字情報の種別がメールアドレスであった場合 にそのメールアドレスへの電子メール作成画面を表示するものである携帯端末装 置。

【請求項14】 請求項11記載の携帯端末装置であって、

前記認識文字活用手段は、前記文字情報の種別がURL (Uniform Resource Locator)であった場合にそのURLへのネットワーク接続画面を表示するものである携帯端末装置。

【請求項15】 請求項4記載の携帯端末装置であって、

前記撮影手段は、認識対象となる文字画像を連続的に撮影する機能を有し、 前記画像取り込み手段は、前記撮影手段により認識対象となる文字画像を複数 取得した場合に、これらの複数の文字画像の画像データを連結して取り込み、

前記文字認識手段は、前記連結された文字画像の画像データの文字認識を行って文字情報に変換するものである携帯端末装置。

【請求項16】 請求項4記載の携帯端末装置であって、

前記文字認識手段は、認識対象の文字種別に応じた複数の認識モードを有し、 予め設定された認識モードによって対応する文字種別に適した文字認識処理を行 うものである携帯端末装置。

【請求項17】 認識対象となる文字画像を文字認識用文字枠とともに表示して撮影するステップと、

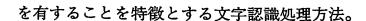
前記撮影した文字画像の画像データを取り込むステップと、

前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するステップと、

前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するステップと、

前記文字配置の解析結果を基に文字画像を切り出すステップと、

前記切り出された文字画像を文字として認識し文字情報に変換するステップと



【請求項18】 請求項17に記載の各ステップをコンピュータに実行させるための文字認識処理プログラム。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】

本発明は、携帯電話機や携帯情報端末等の携帯端末装置に用いて好適な文字認識処理装置及び文字認識処理方法、並びに携帯端末装置に関する。

[0002]

【従来の技術】

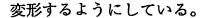
従来の文字認識方法及び装置として、例えば特開2001-60250号公報 (特許文献1)や特開2000-76378号公報(特許文献2)に開示されているものが知られている。

[0003]

特開2001-60250号公報に開示された文字認識方法及び装置は、画像入力された文字から特徴量を抽出し、抽出した特徴量に基づいて文字の認識を行うものである。この従来例では、画像入力された文字を所定間隔で走査する毎に走査線と文字との交差回数を取得するとともに、取得した交差回数の隣接する走査線毎の差分の総和を特徴量の1つとして文字の認識を行うものである。認識対象となる文字に対する走査は、少なくとも2つの異なる方向に所定間隔に行ったり、直交する2方向に所定間隔で行うようにしている。

[0004]

他方、特開2000-76378号公報に開示されている文字認識方法は、入力された画像データの輪郭線データを作成し、作成した輪郭線データを正規化して得たデータを、予め登録している標準文字の輪郭線データを抽出して作成した辞書データに照合して文字認識を行うものである。この際、入力された画像データの書体や歪などによっては、辞書データとの照合前に画像データの輪郭線データを辞書データで認識できる書体に変形したり、あるいは歪が是正されるように



[0005]

【特許文献1】

特開2001-60250号公報

【特許文献2】

特開2000-76378号公報

[0006]

【発明が解決しようとする課題】

しかしながら、従来の文字認識方法及び装置においては、取り込んだ文字画像 データから文字認識処理を行う際に、認識率や認識速度の向上を図るためには、 レイアウト解析や行、文字の切り出し作業においてそれぞれ走査線や輪郭線デー タ等を用いて特別な処理を行う必要があった。

[0007]

また、一般に文字認識を行う文字列全体を一度に撮影する場合、1文字に対する解像度を確保するには撮影装置(例えばCCDカメラなど)の画素数を大きくしなければならない上、必要以上に大規模なデータを扱うために、処理に必要なホストCPUやDSPのMIPS(Million Instructions Per Second)値やワークRAMを大きくしなければならないという問題点もある。

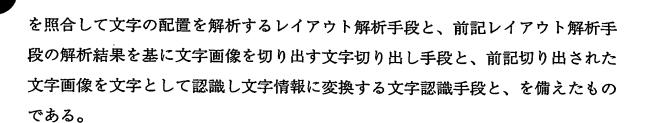
[0008]

本発明は、上記事情に鑑みてなされたもので、その目的は、認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能な文字認識処理装置及び文字認識処理方法、並びに携帯端末装置を提供することにある。

[0009]

【課題を解決するための手段】

本発明に係る文字認識処理装置は、文字認識用に撮影された認識対象となる文字画像の画像データを取り込む画像取り込み手段と、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報



[0010]

上記構成では、文字認識対象物の文字認識を行う際に、認識対象となる文字画 像を文字認識用文字枠とともに液晶表示素子等による表示部に表示して、認識対 象文字列に文字認識用文字枠のカーソルを合わせた状態で文字画像をCCDカメ ラ等の撮像部により撮影する。そして、画像取り込み手段で文字画像の画像デー タを取り込み、レイアウト解析手段により取り込んだ文字画像の画像データに対 してカーソル位置情報を照合して行や文字等の配置を解析し、文字切り出し手段 及び文字認識手段によって文字画像に対する文字の切り出し及び文字認識を行う 。このため、レイアウト解析手段においてカーソル位置情報を利用して認識対象 文字列の行や文字の配置情報を正確に解析することが可能となる。これにより、 認識率及び認識速度の向上が図れる。また、カーソル位置情報を利用することで 文字認識用の画像データの画素数を多くしなくても精度の高い文字認識が可能で あるため、撮影部を多画素化するなどの高性能の仕様は必要なく、装置に要求さ れるスペックの軽減が図れる。さらに、装置内でのデータ処理量を低減すること ができ、これによって、CPUやDSP、メモリ等のシステムを構成するデバイ スに必要とされるスペックを軽減できるため、装置各部の仕様の程度を低減でき 、低消費電力化及びコストダウンを図れる。

[0011]

また、上記構成において、前記画像取り込み手段は、認識対象となる文字画像が連続撮影により得られた複数の文字画像からなる場合に、この文字画像全体から所定領域ずつ文字画像の画像データを分割して取り込み、前記レイアウト解析手段は、前記取り込んだ複数の文字画像の画像データごとに、前記カーソル位置情報を照合して文字の配置を解析するものとする。

[0012]

上記構成では、文字認識対象物の認識対象文字列が大きい場合など、1度の撮



影では文字画像全体を撮影できない場合に、撮影部を移動させながら連続撮影を行って文字画像を撮影する。そして、得られた文字画像の画像データについて、画像取り込み手段で文字画像全体から所定領域ずつ文字画像の画像データを分割して取り込み、レイアウト解析手段により取り込んだ複数の文字画像の画像データごとにカーソル位置情報を照合して行や文字等の配置を解析する。このように文字画像をいくつかのコマの静止画像に分けてカーソル位置情報とともに取り込むことにより、認識対象文字列が長い場合などでもレイアウト解析手段においてカーソル位置情報を利用して認識対象文字列の配置情報を正確に解析することが可能となる。

[0013]

また本発明は、上記いずれかに記載の文字認識処理装置を備えた携帯端末装置を提供する。

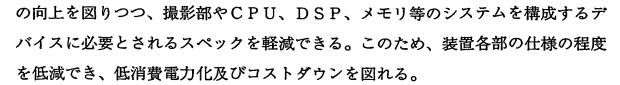
上記構成によれば、文字認識処理装置を搭載した携帯端末装置において、文字 認識の認識率及び認識速度の向上を図りつつ、撮影部やCPU、DSP、メモリ 等のシステムを構成するデバイスに必要とされるスペックを軽減できるため、装 置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

[0014]

本発明に係る携帯端末装置は、文字認識用の認識対象となる文字画像を撮影する撮影手段と、前記撮影された文字画像の画像データを取り込む画像取り込み手段と、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するカーソル情報出力手段と、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するレイアウト解析手段と、前記レイアウト解析手段の解析結果を基に文字画像を切り出す文字切り出し手段と、前記切り出された文字画像を文字として認識し文字情報に変換する文字認識手段と、を備えたものである。

[0015]

上記構成によれば、文字画像の画像データと文字認識用文字枠の位置を示すカーソル位置情報とに基づき、文字認識用文字枠に対応する文字のレイアウト解析を行って文字認識処理を実行することによって、文字認識の認識率及び認識速度



[0016]

また、上記構成において、前記撮影手段による撮影時に、前記文字画像と重ねて文字認識用文字枠を表示する文字枠表示手段を備えるものとする。

この構成によれば、使用者は表示された文字認識用文字枠を見ながら容易に認識対象の文字を文字認識用文字枠に合わせて撮影し、文字認識を行うことが可能となる。

[0017]

また、上記いずれかの構成において、前記文字認識手段による認識結果の文字情報を表示する認識文字表示手段を備えるものとする。

この構成によれば、使用者は認識結果の文字情報を表示によって容易に確認可能となる。

[0018]

また、上記構成において、前記認識文字表示手段は、前記認識結果の文字情報を所定の文字単位ごとに個別に選択可能に表示するものとする。

この構成によれば、使用者は表示された文字情報を基に、認識した文字情報を文字種別などの単位ごとに個別に選択して活用することが可能となる。

[0019]

また、上記いずれかの構成において、前記文字認識手段による認識結果の文字 情報を保存する認識文字保存手段を備えるものとする。

この構成によれば、使用者は認識した文字情報を所望の形態で保存して活用することが可能となる。

[0020]

また、上記構成において、前記認識文字保存手段は、前記文字情報を認識文字 格納領域に保存するものであることとする。

この構成によれば、認識文字格納領域において例えば複数の認識結果の文字情報を格納して利用することが可能となる。



[0021]

また、上記構成において、前記認識文字保存手段は、前記文字情報の種別が電話番号、メールアドレス、URL (Uniform Resource Locator) のいずれかであった場合に、その文字情報の種別に応じたデータベースに登録するものであることとする。

この構成によれば、電話番号、メールアドレス、URLなどを文字認識した場合に、その文字情報の種別に応じて電話帳やブックマークなどのデータベースへ登録して利用することが可能となる。

[0022]

また、上記いずれかの構成において、前記文字認識手段による認識結果の文字 情報を当該文字情報の種別に応じて活用する認識文字活用手段を備えることとす る。

この構成によれば、認識した文字情報の種別に応じて、例えば電話番号、メールアドレス、URLなどの種別毎に文字情報を使用者が有効に利用することが可能となる。

[0023]

また、上記構成において、前記認識文字活用手段は、前記文字情報の種別が電話番号であった場合にその電話番号への発信画面を表示するものとする。

この構成によれば、認識した文字情報が電話番号であった場合にその電話番号 への発信画面を表示し、使用者は簡便に発信動作を行うことが可能となる。

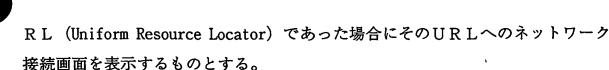
[0024]

また、上記構成において、前記認識文字活用手段は、前記文字情報の種別がメールアドレスであった場合にそのメールアドレスへの電子メール作成画面を表示するものとする。

この構成によれば、認識した文字情報がメールアドレスであった場合にそのメールアドレスへの電子メール作成画面を表示し、使用者は簡便に電子メールの作成、送信動作を行うことが可能となる。

[0025]

また、上記構成において、前記認識文字活用手段は、前記文字情報の種別がU



この構成によれば、認識した文字情報がURLであった場合にそのURLへのネットワーク接続画面を表示し、使用者は簡便にインターネット等のネットワークへの接続動作を行って目的の情報を得ることが可能となる。

[0026]

また、上記いずれかの構成において、前記撮影手段は、認識対象となる文字画像を連続的に撮影する機能を有し、前記画像取り込み手段は、前記撮影手段により認識対象となる文字画像を複数取得した場合に、これらの複数の文字画像の画像データを連結して取り込み、前記文字認識手段は、前記連結された文字画像の画像データの文字認識を行って文字情報に変換するものとする。

この構成によれば、認識対象の文字列が1回の撮影範囲に収まらない場合に、連続的に文字画像を文字認識用文字枠と対応づけて撮影し、得られた複数の文字画像を連結することで、長い文字列であっても文字認識の認識率及び認識速度の向上を図ることが可能である。

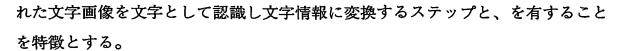
[0027]

また、上記いずれかの構成において、前記文字認識手段は、認識対象の文字種別に応じた複数の認識モードを有し、予め設定された認識モードによって対応する文字種別に適した文字認識処理を行うものとする。

この構成によれば、使用者が文字認識しようとする文字種別に対応した認識モードを設定することで、その文字種別に適した文字認識処理を行うことができ、 認識率を高めることが可能となる。

[0028]

本発明に係る文字認識処理方法は、認識対象となる文字画像を文字認識用文字枠とともに表示して撮影するステップと、前記撮影した文字画像の画像データを取り込むステップと、前記文字画像に対応づけられた文字認識用文字枠の位置を示すカーソル位置情報を出力するステップと、前記取り込んだ文字画像の画像データに対して前記カーソル位置情報を照合して文字の配置を解析するステップと、前記文字配置の解析結果を基に文字画像を切り出すステップと、前記切り出さ



[0029]

この方法によれば、カーソル位置情報を利用して認識対象文字列の行や文字の配置情報を正確に解析することが可能となり、認識率及び認識速度の向上が図れる。また、カーソル位置情報を利用することにより、装置各部に要求されるスペックの軽減が図れるとともに、装置内でのデータ処理量を低減できるため、装置の仕様の程度を低減でき、低消費電力化及びコストダウンを図ることが可能となる。

[0030]

また本発明は、上記の各ステップをコンピュータに実行させるための文字認識 処理プログラムを提供する。

このプログラムによれば、文字画像の画像データと文字認識用文字枠の位置を示すカーソル位置情報とに基づき、文字認識用文字枠に対応する文字のレイアウト解析を正確に行って文字認識処理を実行でき、文字認識の認識率及び認識速度の向上を図りつつ、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

[0031]

【発明の実施の形態】

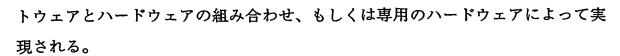
以下、図面を参照して本発明の実施形態を説明する。

(第1実施形態)

図1は、本発明の第1実施形態に係る文字認識処理装置を備えた携帯端末装置 の構成を示すブロック図である。本実施形態では、文字認識処理装置2を携帯端 末装置1に搭載した構成例を説明する。

[0032]

本実施形態の文字認識処理装置 2 は、画像取り込み部 3 と、カーソル制御部 4 と、レイアウト解析部 5 と、文字切り出し部 6 と、文字認識部 7 と、フォーマット出力部 8 とを備えて構成される。これらの構成要素は、プロセッサ及びメモリ等を有してなる制御処理部におけるソフトウェアプログラムの動作、またはソフ



[0033]

携帯端末装置1は、文字認識処理装置2とともに、例えばCCDカメラ等からなる撮影部9と、表示制御部10と、アプリケーション部11とを有している。撮影部9は、文字認識対象物を撮影して撮影データとして取り込み、文字認識処理装置2の画像取り込み部3に入力する。画像取り込み部3は、画像取り込み手段に相当するもので、入力された撮影データを所定のフォーマットの画像データに変換して表示制御部10に入力する。カーソル制御部4は、カーソル情報出力手段に相当するもので、文字認識用文字枠として機能するカーソルに関するカーソル位置情報を表示制御部10に入力する。表示制御部10は、入力された画像データとカーソル位置情報とを表示データに変換し、さらに液晶表示素子等からなるディスプレイ12に適したフォーマットに加工して出力し、ディスプレイ12において情報表示を行う。

[0034]

レイアウト解析部 5 は、レイアウト解析手段に相当するもので、文字認識時に画像取り込み部 3 から所定フォーマットの画像データを受け取るとともに、カーソル制御部 4 からカーソル位置情報を受け取り、カーソル位置情報及び画像データを照合してレイアウト解析を行う。文字切り出し部 6 は、文字切り出し手段に相当するもので、レイアウト解析結果を元に画像データを文字単位で切り出して抽出する。文字認識部 7 は、文字認識手段に相当するもので、切り出された文字単位の画像データの認識処理を行って文字データに変換する。フォーマット出力部 8 は、文字認識部 7 で認識された文字データを適切なフォーマットに変換する。アプリケーション部 1 1 は、フォーマット出力部 8 でフォーマット変換された文字データをアプリケーション用のデータとして利用し、必要に応じて表示制御部 1 0 に入力してディスプレイ 1 2 に表示する。

[0035]

次に、図2を参照して第1実施形態の文字認識処理装置の動作について説明する。図2は第1実施形態に係る文字認識処理時の動作説明図であり、認識対象文



字列を撮影している状態を示している。.

[0036]

携帯端末装置1の撮影部9によって文字認識対象物16の中の認識対象文字列17を撮影し、文字認識を行う。撮影している認識対象文字列17をディスプレイ12に表示する際に、認識対象文字列17と併せて文字認識用文字枠であるカーソル18が表示される。使用者は撮影位置を調整しながら認識対象文字列17をカーソル18に合わせ、この状態で操作ボタン19を操作して静止画像の撮影を実行する。

[0037]

撮影された認識対象文字列17の静止画像は、画像取り込み部3で定められたフォーマットの画像データに変換されて取り込まれ、レイアウト解析部5に渡される。また、レイアウト解析部5ではカーソル制御部4からカーソル位置情報を受け取り、カーソル位置情報と文字画像の画像データとを照合する。このように、レイアウト解析部5では、カーソル位置情報を利用して認識対象文字列17の行及び文字の配置情報を正確に解析することができる。そして、文字切り出し部6で画像データの文字単位での切り出しが行われ、文字認識部7において抽出された文字単位の画像データが文字データに変換され、文字認識が行われる。

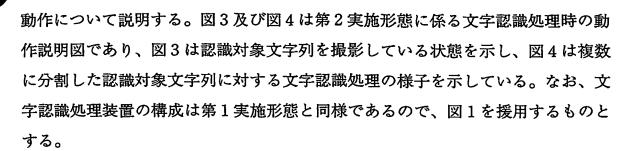
[0038]

このように、第1実施形態の文字認識処理装置では、文字認識対象物16の認識対象文字列17を文字認識する際に、ディスプレイ12に認識対象文字列17の画像と併せて文字認識用文字枠のカーソル18を表示させておき、使用者が認識対象文字列17にカーソル18を合わせた状態で静止画像の撮影を行い、撮影された文字画像の画像データをカーソル位置情報とともに取り込むようにする。これにより、レイアウト解析部5においてカーソル位置情報を利用して行や文字の配置情報を容易かつ正確に解析することができ、文字認識における認識率及び認識速度の向上を図ることができる。

[0039]

(第2実施形態)

次に、図3及び図4を用いて本発明の第2実施形態に係る文字認識処理装置の



[0040]

第2実施形態では、第1実施形態と同様、携帯端末装置1の撮影部9によって 文字認識対象物22の中の認識対象文字列23を撮影して文字認識を行う。この 場合、認識対象文字列23の範囲が広いために1度に取込めない場合は以下のよ うに操作する。

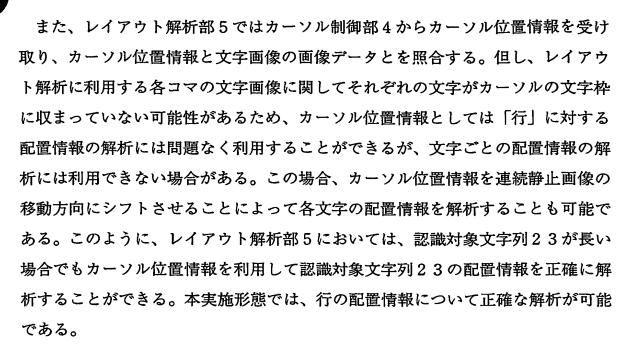
[0041]

撮影している認識対象文字列23をディスプレイ12に表示する際に、認識対象文字列23と併せて文字認識用文字枠であるカーソル18が表示される。使用者は撮影位置を調整しながら認識対象文字列23の開始端位置の部分(図3の例では左端の部分)をカーソル18に合わせ、この状態で操作ボタン19を操作して連続静止画像撮影を開始する。撮影開始後、使用者は携帯端末装置1の移動方向を固定しながら、文字の筆記方向に向かうように撮影開始時の開始端位置と逆方向(図3の例では右方向)に認識対象文字列23に対して平行にスライドさせ、認識対象文字列23の終了端位置(図3の例では右端位置)までカーソル18に収まるようにディスプレイ12に表示させたところで連続静止画像撮影を終了する。このとき、使用者は例えば連続静止画像撮影の最中は操作ボタン19を押下したままにする。

[0042]

撮影された認識対象文字列23の連続静止画像は、画像取り込み部3で図4のように複数の画像データに分割されて取り込まれる。この場合、文字画像全体から1行分の文字画像についていくつかのコマ(1)~(4)の静止画像に分割されて抽出された後、定められたフォーマットの画像データに変換されてレイアウト解析部5に渡される。

[0043]



[0044]

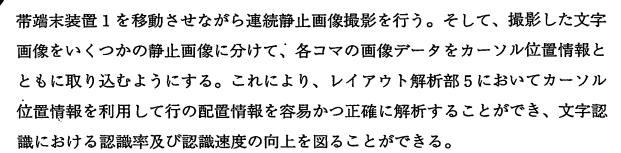
そして、文字切り出し部6で文字単位の切り出しが行われ、文字認識部7において抽出された文字単位の画像データが文字データに変換され、文字認識が行われる。この場合、文字認識結果として、コマ(1)では「ABCDEFGHIJ KL」が得られ、コマ(2)では「FGHIJ KLMNOP」が得られ、コマ(3)では「JKLMNOPQRSTU」が得られ、コマ(4)では「OPQRS TUVW」が得られる。文字認識部7では、これらのコマ毎の文字認識結果を合成することにより、最終文字認識結果として「ABCDEFGHIJ KLMNO PQRSTUVW」が得られる。

[0045]

なお、第2実施形態では、簡単のため1行分の文字画像をカーソルと照合して 一連の文字認識の処理をするかのごとく説明したが、図3に示すようにカーソル 18が2行分存在する場合、一連の文字認識の処理についても2行分同時に行う ことができる。

[0046]

このように、第2実施形態の文字認識処理装置では、文字認識対象物22の認識対象文字列23を文字認識する際に、ディスプレイ12に認識対象文字列17の画像と併せて文字認識用文字枠のカーソル18を表示させておき、使用者が携



[0047]

なお、上記の第2実施形態においては、連続静止画像撮影を行って複数の画像 データを得る際に、携帯端末装置1の操作ボタン19を押下したままで認識対象 文字列上を移動させるようにしているが、これに限るものではない。例えば、使 用者が携帯端末装置1本体を動かして撮影範囲を移動させながらその都度操作ボ タン19を押下することで断続的に複数の画像データを取得し、認識対象文字列 の画像として連続した広い範囲の画像データを得ることも可能である。

[0048]

上述したように本実施形態では、カーソル位置情報を利用することで文字認識用の画像データの画素数を多くしなくても精度の高い文字認識が可能であるため、撮影部9を多画素化するなどの高性能の仕様は必要なく、装置に要求されるスペックの軽減が図れる。さらに、装置内でのデータ処理量を低減することができ、これによって、CPUやDSP、メモリ等のシステムを構成するデバイスに必要とされるスペックを軽減できる。したがって、装置各部の仕様の程度を低減でき、低消費電力化及びコストダウンを図れる。

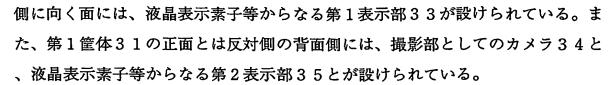
[0049]

(第3実施形態)

次に、第3実施形態として、携帯電話装置に文字認識処理装置を搭載した場合 の構成及び動作の一例を示す。

[0050]

図5は本発明の第3実施形態に係る携帯電話装置の外観構成を示す図であり、 (A) は正面図、(B) は背面図である。本実施形態の携帯電話装置30は、第1筐体31と第2筐体32とを有して構成される。第1筐体31の正面側、すなわち使用者が第1筐体31及び第2筐体32を開いて把持したときに主に使用者



[0051]

第2筐体32の正面側には、オンフック/オフフックキー、数字キー、機能キーなどを含む複数のキーボタンからなる操作部としての操作キー36が配設されている。機能キーPF1~PF5には、装置の動作モードや表示画面などによって各種の機能が割り当てられるようになっている。

[0052]

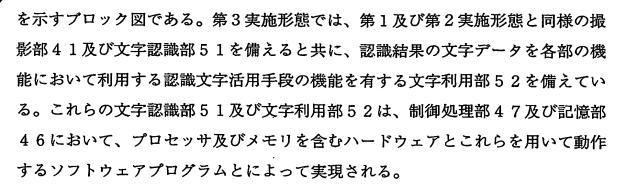
図6は第3実施形態に係る携帯電話装置の構成を示すブロック図である。携帯電話装置は、文字認識対象物等の画像を撮影する撮影手段の機能を有する撮影部41、使用者が操作入力を行う操作部42、文字認識結果等の各種情報を表示する文字枠表示手段及び認識文字表示手段の機能を有する表示部43、移動体通信システムの基地局等との間で無線通信を行う通信部44、音声通話や着信音等の音声処理を行う音声処理部45、各種情報を記憶する認識文字保存手段の機能を有する記憶部46、文字認識処理動作や認識文字データの利用などのアプリケーション処理や各部の制御などを行う制御処理部47を有して構成される。

[0053]

携帯電話装置で音声通話を行う際は、通信部44、音声処理部45及び制御処理部47によって通話音声の入力、送受信、出力を行う。電子メール等のメッセージのやり取りを行ったり、インターネット等のネットワークにアクセスして各種情報を取得する際は、操作部42、表示部43、通信部44及び制御処理部47によって文字データの入力、文字データや画像データ等の情報の送受信、表示を行う。画像撮影を行う際は、撮影部41及び制御処理部47によって被写体を撮影して撮影画像の画像データを取得する。記憶部46は、通信相手の電話番号やメールアドレス等の電話帳登録情報、電子メール等の文字データ、撮影画像等の画像データ、着信音等の音声データなどの各種情報を記憶する。

[0054]

図7は第3実施形態の携帯電話装置における文字認識処理に関する機能的構成



[0055]

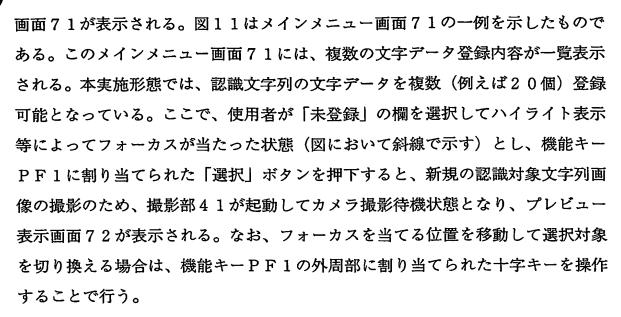
図8は図7の構成における文字利用部の機能的構成を示す説明図である。文字利用部52は、仮保存部61、保存部62、出力部63を有する。仮保存部61は、認識文字保存手段であり、文字認識処理後の文字データを一時記憶するものである。保存部62は、認識文字保存手段であり、認識文字格納領域として複数の保存領域を有し、複数の文字データをそれぞれ格納して保存するものである。出力部63は、文字データの種別を判別する種別判別部64を有し、認識結果の文字データの種別に応じた出力動作を行うものである。文字認識によって得られた文字データが電話番号の場合は、その電話番号への音声発信またはデータベースである電話帳登録に関する動作を行う。また、メールアドレスの場合は、そのメールアドレスへのメール作成または電話帳登録に関する動作を行う。また、URL (Uniform Resource Locator) の場合は、そのURLへアクセスするインターネット接続またはデータベースであるブックマーク登録に関する動作を行う。

[0056]

次に、第3実施形態における文字認識及び文字利用に関する動作の具体例について説明する。図9は文字認識及び文字利用に関する全体の処理の流れを示す説明図である。この図9は、表示部43に表示する表示画面を中心に各動作状態の遷移を示したものである。ここでは、図10に示すような文字認識対象物である名刺101に記載された文字を認識対象として撮影し、携帯電話装置に搭載したOCR (Optical Character Recognition)機能によって文字認識を行う場合を説明する。

[0057]

携帯電話装置のOCR機能を起動すると、初期画面に相当するメインメニュー



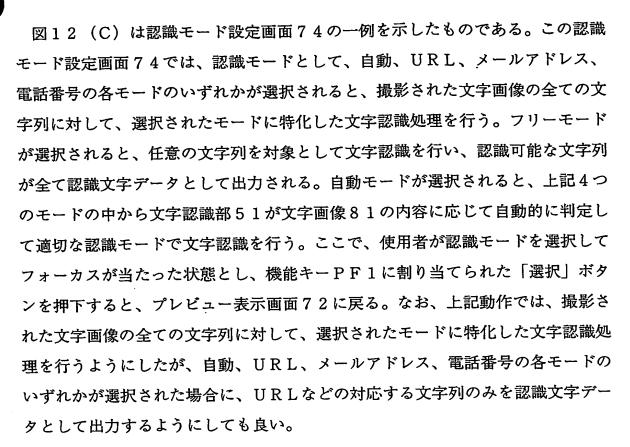
[0058]

図12(A)はプレビュー表示画面72の一例を示したものである。このプレビュー表示画面72には、撮影部41で撮影する認識対象の文字画像81が表示され、この文字画像81上に文字認識領域を示す文字認識用文字枠82が重畳表示される。使用者が文字画像81を見ながら認識対象文字列が文字認識用文字枠82内に収まるように撮影範囲を調整し、機能キーPF1に割り当てられた「撮影」ボタンを押下すると、文字画像81の撮影が行われ、文字認識部51によって認識対象文字列のレイアウト解析、文字切り出し、文字認識等の第1実施形態で述べた一連の文字認識処理が行われる。

[0059]

また、図12(A)のプレビュー表示画面72において使用者が機能キーPF3に割り当てられた「機能」ボタンを押下すると、撮影部41の機能設定モードに移行し、カメラ機能設定画面73が表示される。図12(B)はカメラ機能設定画面73の一例を示したものである。このカメラ機能設定画面73では、設定対象として、認識モード、明暗調整、フォトライト(撮影用照明)のいずれかを選択できるようになっている。ここで、使用者が「認識モード」の欄を選択してフォーカスが当たった状態とし、機能キーPF1に割り当てられた「選択」ボタンを押下すると、認識モード設定画面74が表示される。

[0060]



[0061]

プレビュー表示画面 7 2 の状態から撮影及び文字認識処理が完了すると、認識文字表示画面 7 5 が表示される。図 1 3 (A) は認識文字表示画面 7 5 の一例を示したものである。この認識文字表示画面 7 5 では、認識結果の文字列の文字データを表示する。ここで、認識文字列に電話番号、メールアドレス、URLが含まれる場合は、それぞれに下線が付されたアンカー表示がなされ、選択文字列にフォーカスが当たった状態となる。本実施形態では、認識文字列を解析して電話番号、メールアドレス、URLを識別し、それぞれを個別に選択可能であり、文字列の種別に応じた利用ができるようになっている。

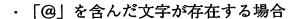
[0062]

電話番号、メールアドレス、URLの識別には、例えば以下の判定方法を用いる。

(1) 電話番号

- ・0から始まる10桁以上の数字が存在する場合
- (2) メールアドレス





(3) URL

・「http://」または「https://」から始まる文字が存在する場合

なお、上記メインメニュー画面 7 1 において、使用者が登録済みの文字列の欄を選択してフォーカスが当たった状態とし、機能キー P F 1 に割り当てられた「選択」ボタンを押下した場合は、認識文字表示画面 7 5 が表示され、過去に認識されて保存された文字列の文字データが表示される。

[0064]

[0063]

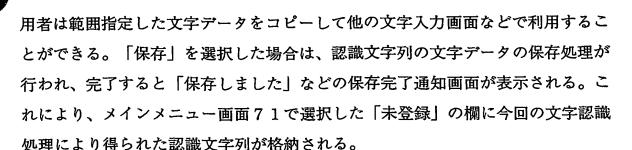
認識文字表示画面 7 5 において使用者が機能キーPF 3 に割り当てられた「機能」ボタンを押下すると、機能メニュー表示画面 7 6 が表示される。図 1 3 (B) は機能メニュー表示画面 7 6 の一例を示したものである。機能メニュー表示画面 7 6 では、連結、編集、コピー、保存、電話帳新規登録、電話帳追加登録、ブックマーク登録のいずれかを選択できるようになっている。

[0065]

一回で撮影できない長い文字列について続けて撮影し、認識結果を一つにつなぎ合わせる場合は、機能メニュー表示画面76で「連結」の機能を選択する。使用者が「連結」の欄を選択してフォーカスが当たった状態とし、機能キーPF1に割り当てられた「選択」ボタンを押下すると、再度プレビュー表示画面72が表示され、撮影可能な状態となる。そして前回に続く認識対象文字列の撮影及び文字認識処理が完了すると、認識文字表示画面75に切り替わり、今回認識された文字列が追加して表示される。

[0066]

機能メニュー表示画面76において使用者が「連結」以外の各機能を選択して機能キーPF1に割り当てられた「選択」ボタンを押下すると、それぞれの機能に対応する機能画面77が表示され、各機能が実行可能となる。「編集」を選択した場合は、テキストエディタが起動して編集画面が表示される。使用者は編集画面において誤認識された文字について修正を行うことができる。「コピー」を選択した場合は、コピー機能が呼び出されてコピー範囲選択画面が表示され、使



[0067]

また、「電話帳新規登録」または「電話帳追加登録」を選択した場合は、電話帳登録画面が表示され、使用者は電話番号やメールアドレス、名前などの認識文字列を電話帳の所望の欄に登録することができる。「ブックマーク登録」を選択した場合は、ブックマーク登録画面が表示され、使用者はURLの認識文字列をブックマークとして登録することができる。

[0068]

認識文字表示画面 7 5 において、使用者が電話番号、メールアドレス、URL のいずれかの文字列を選択した状態で機能キー PF1 に割り当てられた「選択」ボタンを押下すると、それぞれの文字列の種別に応じた利用画面 7 8 が表示される。電話番号を選択した場合は、図13 (C)に示すように利用画面として音声発信画面 7 8 a が表示される。この音声発信画面 7 8 a では、文字認識した電話番号宛の音声発信を簡易な操作で行うことが可能である。URLを選択した場合は、図13 (D)に示すように利用画面としてインターネット接続画面 7 8 b が表示され、文字認識したURLによるインターネット接続を簡易な操作で行うことが可能である。メールアドレスを選択した場合は、図13 (E)に示すように利用画面としてメール作成画面 7 8 c が表示され、文字認識したメールアドレス宛の電子メール作成を簡易な操作で行うことが可能である。

[0069]

このように、第3実施形態では、上記第1及び第2実施形態と同様に、撮影時のプレビュー表示画面において文字認識用文字枠を表示して認識対象の文字画像を撮影し、枠内の画像について文字認識処理を行うことで、文字認識処理にかかる負荷を軽減できる。

[0070]



また、認識モードの選択によって、予め設定した認識対象文字の種別に応じた 文字認識処理を行うことにより、文字認識精度を向上できる。また、認識文字表 示画面において、認識結果として複数の種別の文字データを並べて表示し、所定 の文字単位ごとに分けて個別に選択、利用、保存が可能なようになっている。こ れにより、例えば電話番号の場合は音声発信、メールアドレスの場合はメール作 成など、文字列の種別に応じた認識文字データの利用が容易に可能となる。

[0071]

なお、本発明は上述した実施形態に何ら限定されるものではなく、その要旨を 逸脱しない範囲において種々の態様で実施し得るものである。

[0072]

【発明の効果】

以上説明したように本発明によれば、認識率及び認識速度の向上を図りつつ、 撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減でき、 低消費電力化及びコストダウンを図ることが可能な文字認識処理装置及び文字認 識処理方法を提供できる。

【図面の簡単な説明】

【図1】

本発明の第1実施形態に係る文字認識処理装置を備えた携帯端末装置の構成を 示すブロック図

【図2】

本発明の第1実施形態に係る文字認識処理装置における文字認識処理時の動作 説明図であり、認識対象文字列を撮影している状態を示す図

【図3】

本発明の第2実施形態に係る文字認識処理装置における文字認識処理時の動作 説明図であり、認識対象文字列を撮影している状態を示す図

【図4】

本発明の第2実施形態に係る文字認識処理装置における文字認識処理時の動作 説明図であり、複数に分割した認識対象文字列に対する文字認識処理の様子を示 す図



本発明の第3実施形態に係る携帯電話装置の外観構成を示す図であり、(A) は正面図、(B) は背面図

【図6】

第3実施形態に係る携帯電話装置の構成を示すブロック図

【図7】

第3実施形態の携帯電話装置における文字認識処理に関する機能的構成を示す ブロック図

【図8】

図7の構成における文字利用部の機能的構成を示す説明図

【図9】

第3実施形態における文字認識及び文字利用に関する全体の処理の流れを示す 説明図

【図10】

文字認識対象物の一例を示す図

【図11】

第3実施形態におけるOCR機能のメインメニュー画面の一例を示す図

【図12】

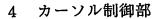
第3実施形態におけるOCR機能の各画面の一例を示したもので、(A)はプレビュー表示画面、(B)はカメラ機能設定画面、(C)は認識モード設定画面をそれぞれ示す図

【図13】

第3実施形態におけるOCR機能の各画面の一例を示したもので、(A)は認識文字表示画面、(B)は機能メニュー表示画面、(C)は音声発信画面、(D)はインターネット接続画面、(E)はメール作成画面をそれぞれ示す図

【符号の説明】

- 1 携带端末装置
- 2 文字認識処理装置
- 3 画像取り込み部



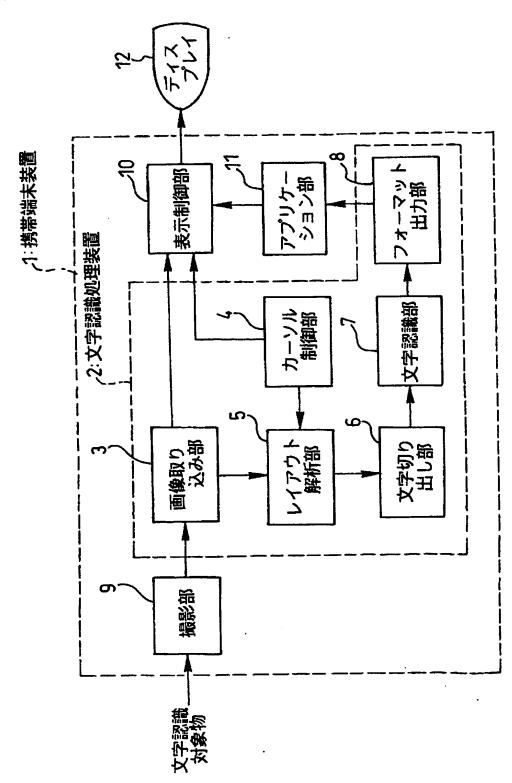
- 5 レイアウト解析部
- 6 文字切り出し部
- 7 文字認識部
- 8 フォーマット出力部
- 9 撮影部
- 10 表示制御部
- 11 アプリケーション部
- 12 ディスプレイ
- 30 携帯電話装置
- 3 3 第 1 表示部
- 34 カメラ
- 3 5 第 2 表示部
- 36 操作キー
- 4 1 撮影部
- 4 2 操作部
- 4 3 表示部
- 4 4 通信部
- 45 音声処理部
- 4 6 記憶部
- 47 制御処理部
- 51 文字認識部
- 52 文字利用部
- 61 仮保存部
- 6 2 保存部
- 6 3 出力部
- 6.4 種別判定部



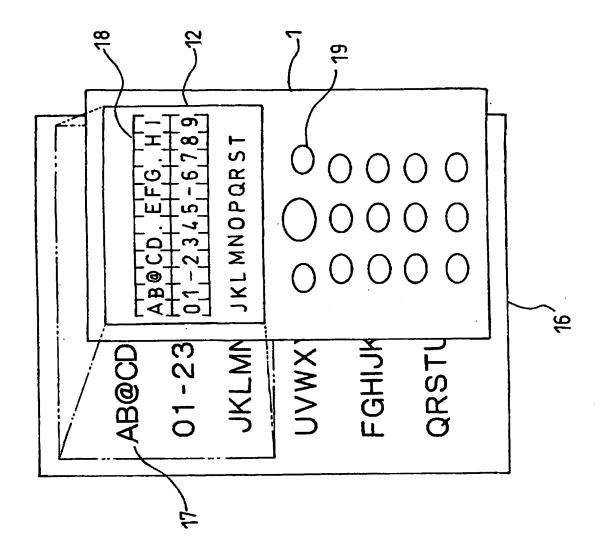
【書類名】

図面

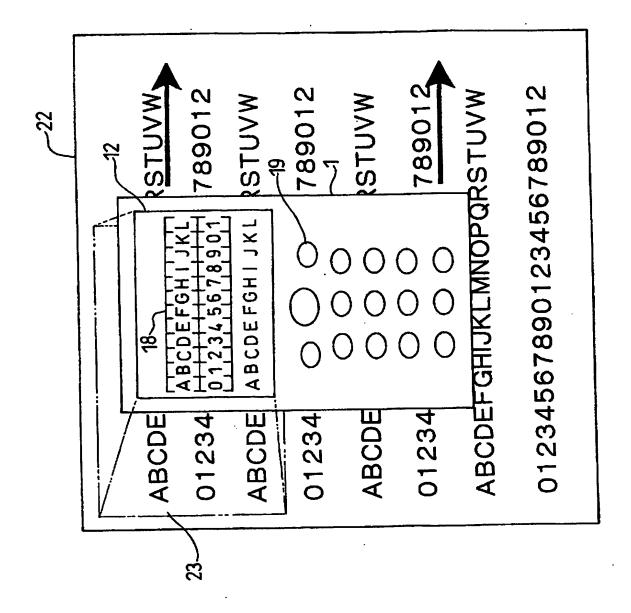
【図1】





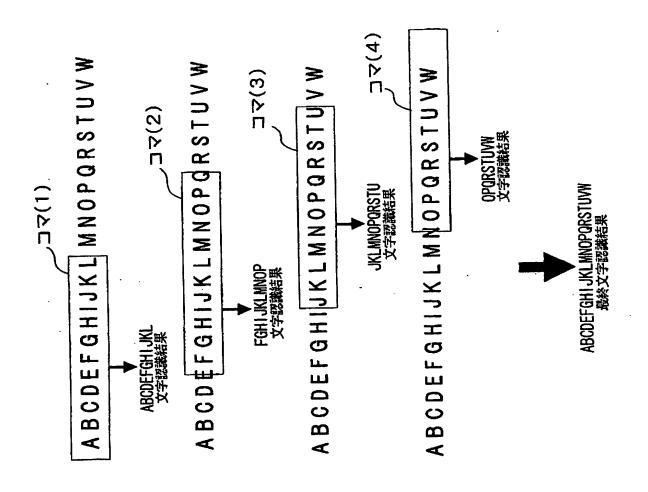






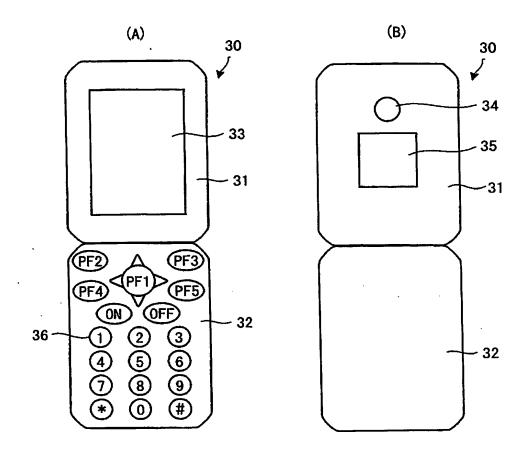


【図4】

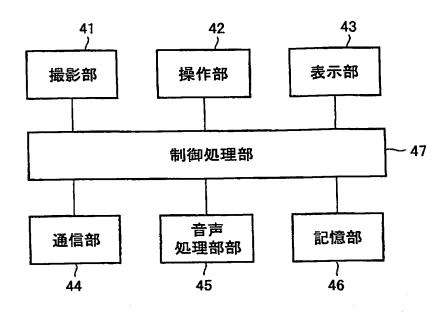




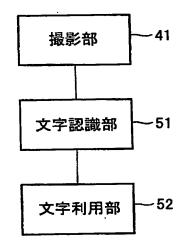




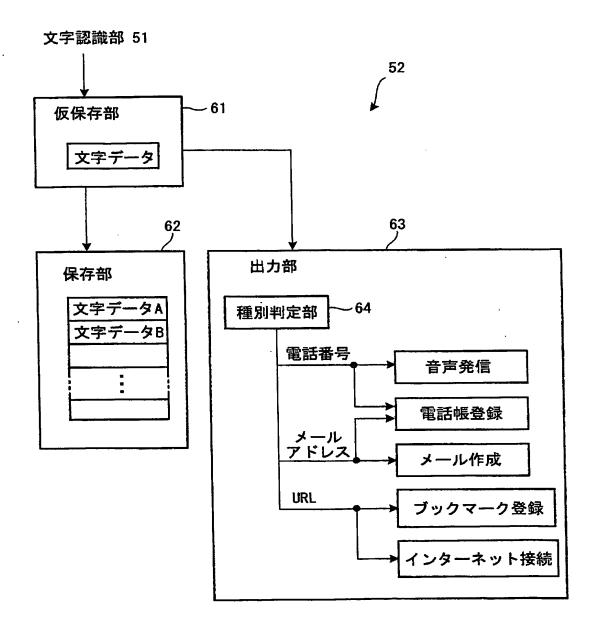




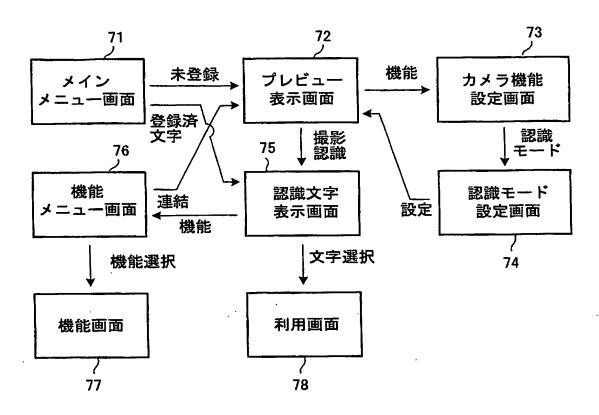
【図7】











【図10】



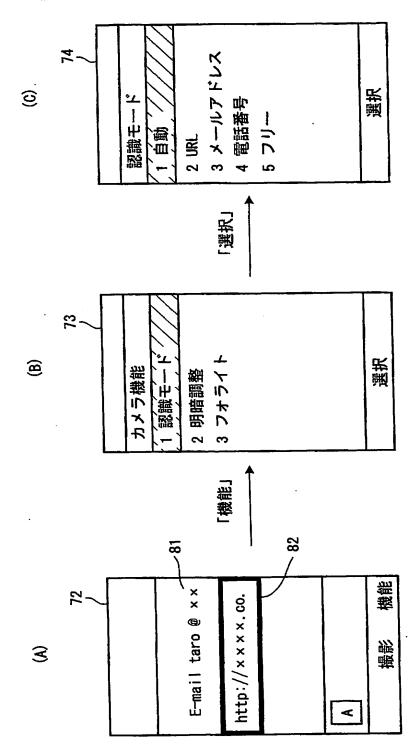


【図11】

71 メインメニュー 1 未登録 2 未登録 3 09012345678 4 http://〇〇〇.com/ 5 未登録 6 未登録 7 未登録 8 未登録

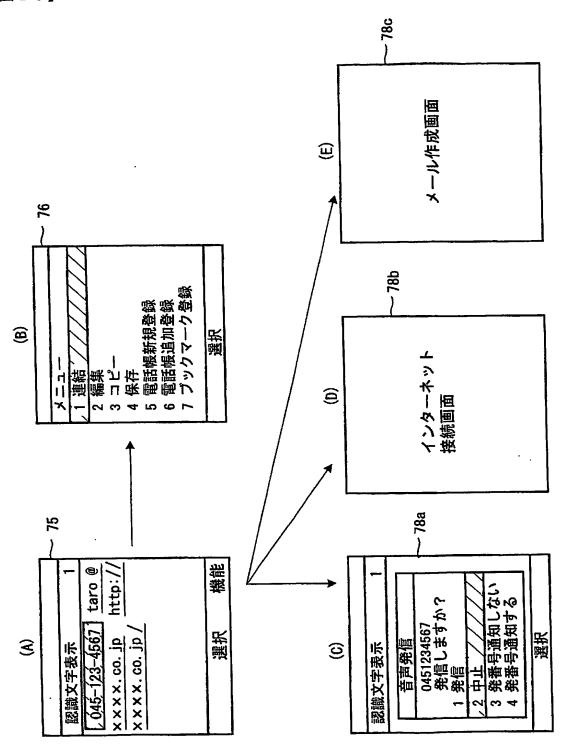








【図13】





【書類名】

要約書

【要約】

【課題】 認識率及び認識速度の向上を図りつつ、撮影装置やシステムを構成するデバイスに必要とされる仕様の程度を低減可能とし、低消費電力化及びコストダウンを図る。

【解決手段】 文字認識対象物の文字画像を文字認識用文字枠のカーソルとともにディスプレイ12に表示し、認識対象文字列にカーソルを合わせた状態で撮影部9により撮影する。画像取り込み部3において、撮影された文字画像の画像データを取り込み、レイアウト解析部5において、取り込んだ文字画像の画像データとこの文字画像に対応するカーソル制御部4からのカーソル位置情報とを受け取り、画像データとカーソル位置情報とを照合して認識対象文字列の行や文字の配置を解析する。そして、文字切り出し部6で解析結果を基に文字の切り出しを行い、文字認識部7によって文字画像に対する文字認識を行って画像データを文字データに変換する。

【選択図】 図1



特願2003-133171

出願人履歴情報

識別番号

[000005821]

1. 変更年月日 [変更理由] 住 所 氏 名

1990年 8月28日 新規登録 大阪府門真市大字門真1006番地 松下電器産業株式会社

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

₩ BLACK BORDERS
IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☑ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
□ other:

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.